COPY __1__ OF __1__ 20-0

HARD COPY $. 100
MICROFICHE $. 0.50

(1)

FUNCTIONAL EQUATIONS IN THE THEORY OF

DYNAMIC PROGRAMMING—VI

A DIRECT CONVERGENCE PROOF

By

Richard Bellman

P-764

8 November 1955

Approved for OTS release

The RAND Corporation

19

Summary

As an approximation to the problem of determining the maximum
of $J(y) = \int_0^T F(x,y)dt$, where x and y are connected by certain
relations, we may consider the question of determining the
maximum of $J(y) = \sum_{k=0}^{N} F(x(k), y(k))$, where $x(k) = x(\frac{kT}{N})$,
$y(k) = y(\frac{kT}{N})$. In this paper we considers the convergence of
the discrete sum as $N \to \infty$, independently of the continuous
process. Convergence is established, using the functional
equation technique of the theory of dynamic programming, under
weaker conditions than those required by the classical calculus
of variations.

FUNCTIONAL EQUATIONS IN THE THEORY OF DYNAMIC PROGRAMMING—VI A DIRECT CONVERGENCE PROOF

By
Richard Bellman

§1. Introduction

A class of variational problems arising frequently in both theory and application involve the maximization of a functional of the form

$$
(1) \qquad J(y) = \int_0^T F(x,y)dt,
$$

subject to relations of the form

$$
(2) \qquad (a) \quad \frac{dx}{dt} = G(x,y), \quad x(0) = c,
$$

$$
(b) \quad R_i(x,y) \leq 0, \quad i = 1, 2, \ldots, K.
$$

Here x and c are N-dimensional vectors, while y is an M-dimensional vector function. The maximization is over y.

Since solutions of problems of this type are only in rare instances obtainable in explicit form, recourse must be had to some type of approximate solution if we are interested in numerical results. A method going back to Euler consists of approximating to the integral in (1) by a sum of the form

$$
(3) \qquad J_1(y) = \sum_{k=0}^{n} \Delta F(x(k), y(k)),
$$

and to the relations in (2) by the relations

(4)   (a) $x(k+1) = x(k) + \Delta G(x(k), y(k)), x(0) = c, k = 0, 1, \ldots, n,$

(b) $R_1(x(k), y(k)) \leq 0,$

where

(5)   (a) $\Delta = T/n,$

(b) $x(k) = x(kT/n), \ y(k) = y(kT/n).$

Under various assumptions concerning the functions F, G and R, it can be shown that

(6) $\lim_{n \to \infty} \ \underset{\{y\}}{Max} \ J_1(y) = \underset{y}{Max} \ J(y),$

or

(7) $\lim_{n \to \infty} \underset{\{y\}}{Max} \ J_1(y) = \underset{y}{Sup} \ J(y).$

Here the maximization on the left—hand side is over all $(n + 1)$—dimensional sequences $[y(0), \ y(1), \ldots, y(N)]$.

An essential feature of the previous proofs of results of this nature is the use of the formulation of the continuous variational process and the existence of a solution. A proof of this type requires relatively strong assumptions concerning the behavior of F and G. In this paper we wish to begin the study of convergence proofs which are independent of the continuous version, and depend only upon the discrete version.

We are interested in imposing relatively weak conditions upon the functions $F(x,y)$ and $G(x,y)$ which will permit us to

conclude that the limit

$$(8) \qquad \lim_{n \to \infty} \operatorname*{Max}_{\{y\}} J_1(y) = J,$$

exists as $n \to \infty$ through some sequence of integers. It is probably true that this limit will **not** exist under the sole assumption of continuity of F and G, together with some simple constraint such as the uniform boundedness of y, although we know of no example of this. We find it necessary to impose Lipschitz conditions of the form $|F(x,y) - F(z,y)| \leq k_1 |x - z|^a$ for all admissible y, where $a^2 + a > 1$. Again we do not know whether this result is best possible.

Apart from the fact that the convergence of the above expression can be established by the methods we present below under lighter conditions than are imposed in the classical calculus of variations, the interest in these results lies in the fact that the techniques we employ open a path to a similar treatment of other types of variational problems.

As we shall see in subsequent papers, eigenvalue problems of one-dimensional and multi-dimensional type, as well as multi-dimensional variational problems of other types, may be treated by these methods.

Furthermore, multi-stage games of continuous type, such as pursuit games, may also be discussed in this fashion, cf. [2]. Here there is no classical theory to guide our analysis, not even a rigorous formulation of the continuous process, nor even the concept of a solution.

In what follows, we shall rely upon the functional equation approach of the theory of dynamic programming, [1], applied to the discrete maximization problem. The results of this paper are self-contained, and require no previous knowledge of either variational theory or dynamic programming.

Finally, let us note in passing, that the method used affords a new approach to the computational solution of variational problems, and a new approach to the determination of analytic properties of the solution such as monotonicity, concavity, and so on.

§2. Functional Equations

In the section following this, we shall discuss various conditions which we can impose upon F, G and R in order to ensure the existence of a maximum in the discrete version. In this section let us proceed upon the assumption that the maximum is attained, and derive the recurrence relations we shall employ to establish convergence.

It is clear to begin with that the maximum of $J_1(y)$, as defined by (1.3), will be a function only of c and n. Let us keep Δ fixed, equal to T/n, and define, for $k = 0, 1, 2, \ldots, n$, the sequence of functions

$$(1) \qquad f_k(c) = \underset{\{y\}}{\text{Max}} \left[\Delta \sum_{j=0}^{k} F(x(j), y(j)) \right],$$

where $x(j)$ and $y(j)$ are subject to the relations

(2)   (a) $x(j+1)=x(j)+\Delta G(x(j),y(j)),x(0)=c,j=0,1,2,\ldots,k,$

(b) $R_1(x(j),y(j))\leq 0,\ k=1,2,\ldots,K,$

and the maximization is now over the finite set of vectors $[y(0),y(1),\ldots,y(k)]$.

Let us now consider this maximization problem as an n-stage decision process, with the state variables c and k. Employing the principle of optimality, cf. [1], it is easy to derive the recurrence relations

(3) $f_{k+1}(c)=\underset{y(0)}{\text{Max}}\ [\Delta P(c,y(0))+f_k(c+\Delta G(c,y(0)))],$

with

(4) $f_0(c) = \underset{y(0)}{\text{Max}}\ \Delta P(c,y(0)).$

The vector $y(0)$ is to be chosen in each case subject to the constraints

(5) $R_1(c,y(0)) \leq 0,\ 1 = 1,\ 2,\ \ldots,\ K.$

The remainder of the convergence proof will be based upon this recurrence relation without further reference to its source.

§3. Conditions Upon P, G and R

Since the method we shall employ is independent of the dimensions of the vectors x and y, there is no loss of generality in taking x and y to be scalar quantities. This will allow us to substitute the usual absolute value notation

for the more distracting norm notation.

We shall assume that $F(x,y)$ and $G(x,y)$ are continuous functions jointly in the variables x and y in some bounded region of the (x,y)-plane. Furthermore, in order to ensure that $x(k)$ does not increase too rapidly, we shall impose a constraint of the form

$$(1) \qquad [G(x,y)] \leq a(y)|x| + b(y),$$

upon $G(x,y)$.

The point of greatest interest is the type of constraint to be imposed upon y. If we remove all constraints of the form $R_1(x,y) \leq 0$, we are in the realm of classical theory. If we impose constraints such as

$$(2) \qquad -\infty < m_1 \leq y \leq m_2 < \infty,$$

or

$$(3) \qquad 0 \leq y \leq x,$$

a constraint which occurs naturally in various classes of multi-stage allocation processes, we enter a region where relatively little has been done in the classical theory apart from the establishment of various types of existence and uniqueness theorems. Since there are greater analytic difficulties in the way of analytic solutions in the case where constraints are present, which means a greater dependence upon numerical techniques, we shall concentrate upon this case.

If we set $y = vx$, we see that (3) becomes

$$(4) \qquad 0 \leq v \leq 1,$$

a constraint of the type given in (2).

Finally we have to impose some conditions on the modulus of continuity of the functions $F(x,y)$ and $G(x,y)$. The simplest is a uniform Lipschitz condition of the type

$$(5) \qquad |F(x_1,y) - F(x_2,y)| \leq k|x_1 - x_2|^a,$$

for some $k > 0$ and a satisfying $0 < a \leq 1$, for all y satisfying (2) and x_1, x_2 lying in some fixed interval $[-c,c]$. For our convergence proof we shall impose a further condition upon a, given below in (1) of §4.

§4. Statement of the Principal Result

The main result we shall establish is

Theorem. Assume that

(1) (a) y satisfies a constraint $-\infty < m_1 \leq y \leq m_2 < \infty$.

    (b) $F(x,y)$ and $G(x,y)$ are jointly continuous in x and y in a region of the form $-c_1 \leq x \leq c_1$, $m_1 \leq y \leq m_2$, and satisfy (3.5) in this region, for an a satisfying $a^2 + a > 1$.

    (c) $|G(x,y)| \leq a_1|x| + b_1$ in this region.

Under these conditions, the sequence $\{f_n(c)\}$ converges to a function $f(c,T)$ as $n \to \infty$ through the sequence of values $\{2^k\}$, uniformly in an interval $[-c_2,c_2] \subset [-c_1,c_1]$, for T sufficiently small, dependent upon the constants appearing above.

The positive solution of $a^2 + a - 1$ is $a = (-1 + \sqrt{5})/2 \simeq .62$.

§5. **A Lemma**

We shall repeatedly use the following result.

Lemma 1. <u>Let</u> $a(c,v)$, $a'(c,v)$, $f(c,v)$ <u>and</u> $f'(c,v)$ <u>be continuous</u>
<u>functions of</u> v <u>for</u> c <u>and</u> v <u>in some region</u> D. <u>Define</u>

$$(1) \qquad F(c) = \underset{v \leq D}{\text{Max}} \left[a(c,v) + f(c,v) \right]$$

$$G(c') = \underset{v' \leq D}{\text{Max}} \left[a'(c',v') + f'(c',v') \right].$$

<u>Then</u>

$$(2) \qquad |F(c) - G(c')| \leq \underset{v \leq D}{\text{Max}} \left[|a(c,v) - a'(c',v')| + |f(c,v) - f'(c',v)| \right].$$

<u>Proof</u>: Let $v = v(c)$ be a value of v which yields the maximum
in the expression for $F(c)$, and $v' = v'(c')$, the corresponding
expression for v'. Then

$$(3) \qquad F(c) = a(c,v) + f(c,v) \geq a(c,v') + f(c,v'),$$

$$G(c') = a'(c',v') + f'(c',v') \geq a'(c',v) + f'(c',v).$$

From these inequalities we derive the further inequalities

$$(4) \qquad F(c) - G(c') \geq a(c,v') - a'(c',v') + f(c,v') - f'(c',v')$$

$$\leq a(c,v) - a'(c',v) + f(c,v) - f'(c',v),$$

from which we obtain

(5) $\qquad |F(c)-G(c')| \leq \text{Max} [|a(c,v')-a'(c',v')|+|f(c,v')-f'(c',v')|,$

$$|a(c,v)-a'(c',v)|+|f(c,v)-f'(c',v)|]$$

This, in turn, yields the inequality in (2).

§6. Uniform Lipschitz Conditions

We start out by assuming that $F(x,y)$ and $G(x,y)$ satisfy Lipschitz conditions in an interval $[-c_1,c_1]$. Starting with c in an interval $[-c_2,c_2]$ at the beginning of an n-stage process, we see that after one stage, we have an x-value lying in the interval $-c_2-\Delta \text{Max } G(x,y) \leq x \leq c_2 + \Delta \text{Max } G(x,y)$, where the maximum may be taken over the fixed interval $-c_1 \leq x \leq c_1$, $m_1 \leq y \leq v_2$. Continuing in this way, we see that the range of x may increase with each stage of the process. This is why we must start in an interval $[-c_2,c_2] \subset [-c_1,c_1]$, and why T may be constrained.

Let $b = \text{Max } G(x,y)$, and let us call the interval $[-c_1 - bk\Delta, c_1 + bk\Delta]$ the kth interval. We choose T and Δ initially so that with c in the initial interval, the nth interval is contained in $[-c_2,c_2]$. In this way we preserve uniform bounds.

In some cases, where $x(t)$ is decreasing as a function of t; i.e., where $G(x,y) \leq 0$ for all x and y, we do not meet this difficulty.

It is essential for our proof to establish a uniform Lipschitz condition for the members of the sequence $\{f_k(c)\}$.

Lemma 2. Consider the sequence $\{f_k(c)\}$ as defined by (2.3) and

(2.4), <u>under the conditions of the theorem of</u> §4. <u>For k =</u>
0, 1, ..., n, <u>we have</u>

(1) $\qquad |f_k(u) - f_k(w)| \leq m|u - w|^a,$

<u>for</u> u <u>and</u> w <u>in the</u> $n - k + 1)^{st}$ <u>interval, where</u> m <u>is independent</u>
<u>of</u> u, w, k <u>or</u> Δ.

<u>Proof</u>: The proof will proceed by an induction on k. We have

(2) $\qquad f_0(u) = \underset{y}{\text{Max}}\ \Delta P(u,y),$

$\qquad\qquad f_0(w) = \underset{y'}{\text{Max}}\ \Delta P(w,y'),$

where y and y' satisfy the constraints $m_1 \leq y, y' \leq m_2$.

Applying Lemma 1, we obtain the inequality

(3) $\qquad |f_0(u) - f_0(w)| \leq \underset{y}{\text{Max}}\ \Delta |P(u,y) - P(w,y)|$

$\qquad\qquad\qquad\qquad \leq K|u - w|^a.$

Assume that we have demonstrated that

(4) $\qquad |f_k(u) - f_k(w)| \leq K_k \Delta |u - w|^a,$

for k = 0, 1, 2, ...,L, for u and w in the $(n - k - 1)^{st}$
interval. Turning to the recurrence relation (2.3), and applying
Lemma 1, we obtain the relation

(5) $\qquad |f_{k+1}(u)-f_{k+1}(w)| \leq \underset{y}{\text{Max}}[\Delta|P(u,y)-P(w,y)|$

$\qquad\qquad\qquad\qquad + |f_k(u+\Delta G(u,y))-f_k(w+\Delta G(w,y))|].$

If u and w lie in the $(n - k - 2)^{nd}$ interval, the points
$u + \Delta G(u,y)$, $v + \Delta G(w,y)$ will certainly be included in the
$(n - k - 1)^{st}$ interval.

$$(6) \qquad |f_{k+1}(u) - f_{k+1}(w)| \leq \Delta K |u-w|^a + K_k \Delta |u+\Delta G(u,y)-w-\Delta G(w,y)|^a.$$

$$\leq (K_k + K)\Delta |u-w|^a + a_1 K_k \Delta^2 |u-w|^a,$$

for a fixed constant a_1.

This shows that we can take $K_k = a_2 kK$, for some constant
$a_2 > 1$. Since $k\Delta \leq n\Delta = T$, we see that we have a uniform
Lipschitz condition.

§7. Stability.

We now wish to demonstrate a result concerning the sta-
bility of the sequence $\{f_k(c)\}$ under perturbations of the function
$F(x,y)$.

Lemma 3. Consider the two sequences

$$(1) \qquad F_{k+1}(c) = \underset{v}{\text{Max}} \left[\Delta F(c,v) + f_k(c + \Delta G(c,v)) \right]$$

$$F_{k+1}(c) = \underset{v}{\text{Max}} \left[\Delta F'(c,v) + F_k(c + \Delta G(c,v)) \right],$$

with

$$(2) \qquad f_0(c) = \underset{v}{\text{Max}} \, \Delta F(c,v),$$

$$F_0(c) = \underset{v}{\text{Max}} \, \Delta F'(c,v).$$

We have, under the hypotheses of the theorem

(3) $$|F_k(c) - f_k(c)| \leq k\Delta \, \underset{v}{\text{Max}} \, \underset{c_k}{\text{Max}} \, |F(c,v) - f'(c,v)|,$$

for $k = 0, 1, 2, \ldots, n$. The notation $\underset{c_k}{\text{Max}}$ signifies that the maximum is taken over the kth interval, as defined above.

Proof: Let us proceed inductively. The result is clearly true for $k = 0$. Assume that we have

(4) $$|F_k(c) - f_k(c)| \leq L_k\Delta \, \underset{v}{\text{Max}} \, \underset{c_k}{\text{Max}} \, |F(c,v) - F'(c,v)|,$$

for $k = 0, 1, 2, \ldots, K$, and let us determine the form of L_{K+1}. Applying Lemma 1, we obtain the result

(5) $$|F_{k+1}(c) - f_k(c)| \leq \underset{v}{\text{Max}} \, [\Delta|F(c,v) - F'(c,v)| + |F_k(c+\Delta G) - f_k(c+\Delta G)|]$$

$$\leq \underset{v}{\text{Max}} \, [\Delta|F(c,v) - F'(c,v)| + L_k\Delta \, \underset{v}{\text{Max}} \, \underset{c_{k+1}}{\text{Max}} \, |F(c,v-F'(c,v)|]$$

$$\leq (L_k + 1)\Delta \, \underset{v}{\text{Max}} \, \underset{c_{k+1}}{\text{Max}} \, |F(c,v) - F'(c,v)|.$$

Hence we may take $L_{k+1} = L_k + 1$.

§8. Transplantation

The next step of the proof depends upon an idea which is abstractly identical with one used by Polya and Schiffer in a similar situation, and called by them "transplantation".

The basic idea is the following. Consider two multi-stage decision processes, having, in general, distinct optimal policies. Inequalities connecting the return functions of the two processes may be obtained by interchanging the role of the optimal policies, using each in the other process.

Sometimes, however, this procedure cannot be carried out, since an optimal policy for one process need not be an admissible policy for the other. Although something of this sort occurs below, we can circumvent the difficulty it causes by using an interesting lemma concerning the convergence of sequences.

We shall employ the idea sketched above in the following way in order to compare the process with intervals, or stages, of length Δ with the corresponding process where the interval is of length 2Δ. Consider the Δ—process under the additional restriction that the choice of $y(2k)$ must be the same as the choice of $y(2k + 1)$. It is clear that the return, i.e., value of the maximum, that we obtain from this process will be less than the return from the original Δ—process. On the other hand, the return from this process should be, granted the principle of wishful thinking, close to the return from the 2Δ—process, for Δ small.

Combining these results, we shall obtain an inequality connecting the returns of the Δ— and 2Δ—processes. This inequality is strong enough to yield convergence.

§9. Description of the Δ— and 2Δ—process.

Let us now make the above remarks precise. To define the 2Δ—process, the interval $[0,T]$ is divided into equal intervals of length 2Δ. Choices of y are made at the points 0, 2Δ, 4Δ, and so on.

The Δ—process is defined similarly, with intervals of

length Δ. Let $\{f_k(c)\}$ denote the sequence of returns from the
Δ-process, as defined by the recurrence relations of (2.3)
and (2.4), and let $\{g_k(c)\}$ denote the sequence of returns from
the 2Δ-process.

Let us now define the following intermediate process.
The interval length is Δ, but the policies are restricted to
those which employ the same y-value at the points $2k\Delta$ and
$(2k+1)\Delta$. Let $\{h_{2k}(c)\}$ denote the sequence of returns obtained
in this way. Then

(1)
$$h_0(c) = \underset{y}{\text{Max}} \ \Delta P(c,y),$$

$$h_2(c) = \underset{y}{\text{Max}} \ [\Delta P(c,y) + \Delta P(c + \Delta G(c,y),y)$$
$$+ h_0(c + \Delta G(c,y) + \Delta G(c + \Delta G(c,y),y)],$$

$$\vdots$$

$$h_{2k+2}(c) = \underset{y}{\text{Max}} \ [\Delta P(c,y) + \Delta P(c + \Delta G(c,y),y)$$
$$+ h_{2k}(c + \Delta G(c,y) + \Delta G(c + \Delta G(c,y),y)].$$

Here y is subject to the constraint $m_1 \leq y \leq m_2$.

It is clear that

(2)
$$h_{2k}(c) \leq f_{2k}(c), \quad k = 0, 1, 2, \ldots$$

Let us now compare $h_{2k}(c)$ with $g_k(c)$.

It is easy to show that the sequence $\{h_{2k}(c)\}$ satisfies
the same type of uniform Lipschitz condition as that which we
derived for $\{f_k(c)\}$. Hence we may write

(3) $$h_{2k+2}(c) = \underset{y}{\text{Max}} \; [2\Delta F(c,y) + h_{2k}(c + 2\Delta G(c,y)) + E_k(c,y)],$$

where

(4) $$|E_k(c,y)| \le a_2 \, \Delta^{a(1+a)},$$

since

(5) $$h_{2k}(c+\Delta G(c,y)+\Delta G(c+\Delta G(c,y),y))=h_{2k}(c+2\Delta G(c,y)+O(\Delta^{1+a}))$$

$$=h_{2k}(c+2\Delta G(c,y))+O(\Delta^{a(1+a)}).$$

Applying Lemma 3, we see that

(6) $$|h_{2k}(c)-g_k(c)| \le a_2(n\Delta)\Delta^{a(1+a)-1} \le a_2 T \, \Delta^b,$$

where $b = a(1 + a) - 1 > 0$.

Combining (2) and (6), we obtain

(7) $$f_{2k}(c) - g_k(c) \ge -a_2 T \Delta^b$$

Now let $\Delta \to 0$ through a sequence $\Delta, \Delta/2, \ldots, \Delta/2^r, \ldots$
Let the return from the kth stage of the Δ-process be $u_1(c)$, the
return from the $2k^{th}$ stage of the $\frac{\Delta}{2}$-process be $u_2(c)$, and,
generally, the return from the $2^r k^{th}$ stage of the $\frac{\Delta}{2^r}$-process
be $u_r(c)$.

From the inequality in (6) we conclude that

(8) $$u_{r+1}(c) \ge u_r(c) - a_2 T(\Delta/2^r)^b.$$

To complete the proof we require a result concerning the
convergence of sequences, which we will prove in the next section.

§10. A Result on Convergence

Let us establish the following result

Lemma 4. **Let** $\{a_n\}$ **be a sequence satisfying the following conditions:**

(1)   (a) $\infty > M \geq a_n \geq a_n - b_n$,

   (b) $b_n \geq 0$, $\Sigma b_n < \infty$.

Then the sequence converges.

Proof: It is clear that the sequence is uniformly bounded from below. Let x_1 and x_2 be two distinct cluster points, with u_M converging to x_1, and u_N converging to x_2. Let $M_1 < N_1 < M_2$, with u_{M_1}, u_{M_2} close to x_1, and u_{N_2} close to x_2. Then, on one hand,

$$(2) \qquad u_{N_1} - u_{M_1} \geq -\sum_{k=M_1}^{N_1} a_k \geq -\epsilon,$$

and on the other hand,

$$(3) \qquad u_{M_2} - u_{N_1} \geq -\sum_{k=N_1}^{M_2} a_k \geq -\epsilon.$$

Since ϵ can be made arbitrarily small, we have $x_1 = x_2$.

§11 Conclusion of the Proof

Applying Lemma 4, we see that $u_r(c)$ converges as $r \to \infty$ to a function $f(c,s)$, where $s = kT/n$. Taking n sufficiently large, we obtain approximations to a continuous function of s, $f(c,s)$.

References

1. Bellman, R., "The Theory of Dynamic Programming", Bull.
   Amer. Math. Soc., Vol. 60 (1954), pp. 503-16.

2. _____ , "Functional Equations in the Theory of
   Dynamic Programming—III, Multi-stage Games,
   (to appear).

BLANK PAGE